Local Associated Features for Pedestrian Detection

Song Shao[1,2], Hong Liu[1,2], Xiangdong Wang[1,2], Yueliang Qian[1,2]

[1]Research Center for Pervasive Computing, Institute of Computing Technology Chinese Academy of Sciences, Beijing 100190, China
[2]Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology Chinese Academy of Sciences, Beijing 100190, China

Abstract. Local features are usually used to describe pedestrian appearance. While most of existing pedestrian detection methods don't make full use of context cues, such as associated relationships between local different locations. This paper proposes two novel kinds of local associated features, gradient orientation associated feature (GOAF) and local difference of ACF (ACF-LD), to exploit context information. In our work, pedestrian samples are enlarged to contain some background regions besides human body, and GOAF, ACF and ACF-LD are combined together to describe pedestrian sample. GOAF are constructed by encoding gradient orientation features from two different positions into a single value. These two positions are come from different distance and different direction. For ACF-LD, the sample is divided into several sub regions and the ACF difference matrixes between these areas are computed to exploit the associated information between pedestrian and surrounding background. The proposed local associated features can provide complementary information for detection tasks. Finally, these features are fused with ACF to form candidate feature pool, and AdaBoost is used to select features and train a cascaded classifier of depth-two decision trees. Experimental results on two public datasets show that the proposed framework can achieve promising results compared with the state of the arts.

1 Introduction

Pedestrian detection is one important task in computer vision and pattern recognition, which detects pedestrian's location and size in images or videos. It has wide applications, such as intelligent surveillance, driver assistance and human behavior analysis.

Currently the predominant approaches of pedestrian detection are based on machine learning. The key factors of these methods are feature representation and classifier construction. Most early methods used single features, such as Haar [1], Local Binary Pattern (LBP) [2], Edgelet [3], Shapelet [4] and Histograms of Oriented Gradients (HOG) [5]. Since the description ability of single feature is limited, multiple features appeared in recent years. For example, Wang [6] connected HOG with LBP features, Wojek [7] combined Haar, HOG and Shapelet features, Walk [8] added self-similar color features and motion features on the basis of Wojek [7]. Dollar [9] proposed integral channel features (ICF), and gave an optimized channel combination including gradient magnitude, histogram of oriented gradients, and LUV colors. These channel features improved the performance of pedestrian detection and can be computed quickly with integral images technology. Dollar [10] further proposed aggregated channel features (ACF), which uses pixel lookups in aggregated channels to reduce feature extraction time without constructing integral images. Each of the ACF feature represent a local block of the sample, which is simpler and computed more quickly. Effective combination of feature channels and feature selecting strategy by cascaded classifiers make ACF framework perform better than most of other methods both on detection accuracy and speed [10].

However, most above methods usually use features in pedestrian region. Although pedestrian appearance contains abundant information for detection, while in complex scenes, such as occlusion, crowded scenes or poor image resolution, it's hard to detect pedestrian effectively. Context information, such as background region around human body has not been fully used. In many dynamic scenes, such as on-board videos, though the backgrounds are changing as time goes on, the structures of scenes are relatively stable, which contains some useful context information. The associated features between different locations in background and pedestrian region are not fully exploited, which would contain more information than single local features just in human regions. We call these features local associated features in this paper. According to the retrieval reference, up to date, ACF is one of the most successful pedestrian descriptors both in detection accuracy and detection speed [10]. While each ACF feature is one-dimensional channel feature and cannot describe the associated information of different local regions, such as the gradient orientation. Besides, ACF is focus on pedestrian area and the difference between pedestrian and surrounding background is not paid enough attention. Therefore, we try to design some local associated features to exploit the context information between different locations, and fuse with ACF to improve the robustness of feature descriptor.

This paper proposes two novel kinds of local associated features for pedestrian detection: Gradient Orientation Associated Feature (GOAF) and Local Difference of ACF (ACF-LD). In our work, pedestrian samples are enlarged to contain some background regions, and GOAF, ACF and ACF-LD are combined together to describe pedestrian sample. Firstly, GOAF associates gradient orientations of different regions in a certain distance and encodes them into a single value, which can exploit associated information between local regions. Besides, to exploit the associated information between pedestrian and surrounding background, object samples are divided into several sub regions and local difference of ACF are calculated between these regions. The proposed local associated features provide complementary context information for pedestrian detection. In our pedestrian detection framework, ACF and proposed local associated features GOAF and ACF-LD are fused to form candidate feature pool. Then AdaBoost is used to select features and train a cascaded classifier of depth-two decision trees. We evaluate our framework on two different pubic datasets. The experimental results show the effectiveness of our method. For example, the miss rate is reduced from 44.04% to 38.07% at 10^{-1} FPPI on Caltech dataset and from 16.85% to 16.04% at 10^{-1} FPPI on Inria dataset. This means that the proposed local associated features can effectively capture context information of scenes and improve pedestrian detection performance compared with the state of the arts.

2 Related Work

Recent studies show that context information plays an important role in video and image understanding. Researchers have proposed many different types of context information, such as semantic context [11], 3D geometry context [12], local pixels context [5, 13] and shape context [14].

Dalal [5] slightly enlarged the detection window to include neighbor pixels around the pedestrian and then extracted HOG features on the enlarged windows. But simply expanding the window just got limited improvement. If the window continues to enlarge, the dimension of features will increase significantly and detection performance will be worse. Neil [15] segmented the image into different kinds of regions such as grass, roads and sky, and then learned the probability of a person appearance in a certain region to adjust the detection results. This approach uses scene context information and the performance relies on image segmentation and probability learn-ing, which is not very effective for complex and dynamic scenes. William [16] pro-posed a feature descriptor called Local Response Context (LRC). This method firstly sampled the detection responses around each detection window to construct a feature vector and then learned a partial least squares regression model as a second classification stage. LRC descriptors reduced the dimensionality of features, but the improvement of performance is limited. Above methods indeed used some context information, while these methods did not exploit the associated relationship between local regions in the same detection window.

Ding [17] computed multi-scale HOG features and put forward a new feature called Local Difference Pattern (LDP), which is similar to LBP. The local region is divided into blocks, and the difference between the average pixel intensity in each block and a reference block forms the LDP feature. In addition, for each detection result, the classifier responses from neighbor locations and scales are incorporated as additional features to join an iterative training process called Context Boost. While the multi-scale features take better advantage of context information, it requires that the pedestrian area have to be half of the detection window. Besides, the whole train-ing and detecting framework is complicated. In our paper, the proposed ACF-LD also adopts region partition method, but the partition schemes are totally different and the segmented sub areas are used to compute the local ACF difference matrixes instead of multi-scale HOGs. Besides, LDP only describe color context cues in each local region, while GOAF connects the gradient orientation values of different local regions.

This paper focuses on constructing novel local associated features, which can provide a new thought for exploiting context information of local features in different regions. The detailed construction of features is introduced in section 3. Section 4 introduces the experiments and discussion. Section 5 summarizes this paper and gives future research work.

3 Our Proposed Method

3.1 The Framework

This paper introduces a pedestrian detection framework based on local associated features and Aggregated Channel Features (ACF) to describing pedestrian object robustly. To exploit and use context information on local features in different regions, we propose two novel kinds of local associated features: Gradient Orientation Associated Feature (GOAF) and Local Difference of ACF (ACF-LD). Then GOAF, ACF and ACF-LD are fused together to form candidate feature pool. Finally, AdaBoost is used to select distinguish features and a cascaded classifier of depth-two decision trees is trained to select and combine these distinguish local features. The whole process of our detection framework is shown in Fig.1.

Pedestrian samples are enlarged to contain some neighbor background regions. The details of enlarge ratio will discussed in next section. In training stage, GOAP, ACF and ACF-LD are extracted from samples and then a multiple round of bootstrapping is used to train decision trees over these candidate features. Finally, all the weak classifiers are cascaded to form the final detector. In testing process, feature pyramids are constructed for every detection image and then a sliding window is used over multiple scales. The features corresponding to the current detection window are sent to the detector for classification. Then the final results are obtained after non-maximal suppression (NMS).

Our main work is the construction of local associated features. The details about sample enlarging strategy, and constructions of GOAF and ACF-LD will be introduced in next sections.

3.2 Gradient Orientation Association Feature (GOAF)

For machine learning based pedestrian detection method, feature descriptor is the most important key for classification and detection performance. In existing features, up to now, HOG [5] and ACF [10] are useful descriptors and over half of ACF comes from HOG channel features. These channel features describe local gradient magnitude in different orientations and effectively describe the pedestrian characteristics.

We enlarge the sample size to include more nearby background areas as Fig.2 shows, similar to Dalal's [5] and Dollar's [10] methods. The parameters of pedestrian size $ph \times pw$ and sample size $sh \times sw$ are determined through experiments and discussed in section 4. The background regions can provide

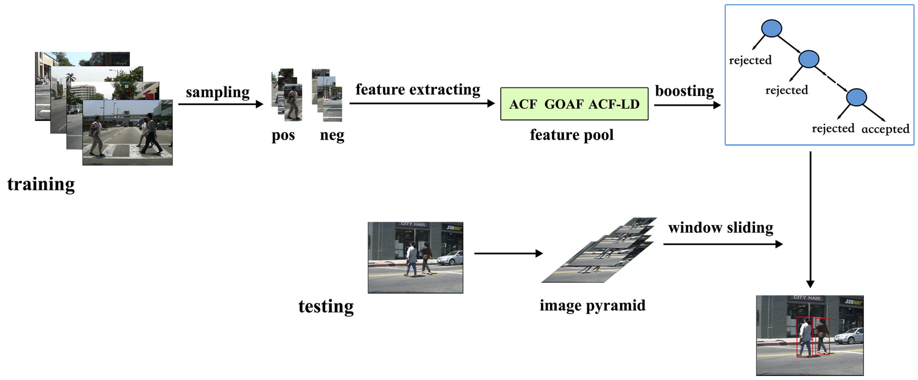

Fig. 1. Our pedestrian detection framework.

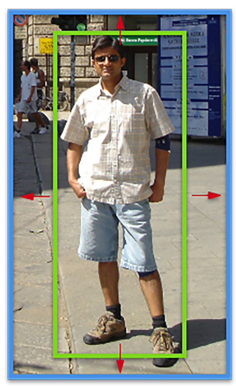

Fig. 2. Enlarging samples to contain some neighbor background regions.

useful context cues. But just expanding the sample is not enough and some strategy should be used to exploit hidden information.

After observed lots of images, we found that there exists some association between the gradient orientations of different human body parts. For example, such as the head, the hands and the legs, even the neighbor background regions and human body parts as Fig.3 shows. This associated information would contain more description abilities than the single local gradient feature.

This paper constructs two local gradient orientation features within certain distance and certain direction in the sample, and encoding them into a single value. This value reflects the gradient information of the two positions at the same time and embodies the association of different local areas. First of all, we compute gradient orientation channel for the sample. In order to calculate quickly, we shrink the sample from $h \times w$ pixels to $(h/shrink) \times (w/shrink)$ pixels, and then extract gradient orientation matrix G. Each value of this matrix is between 0 and π To reduce calculation, these values are normalized averagely

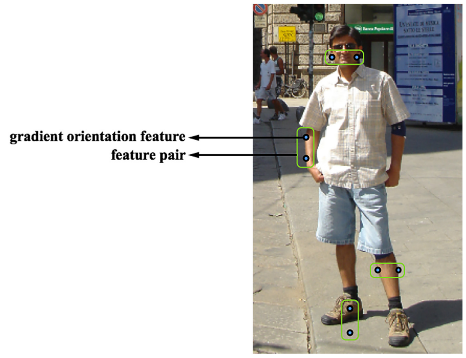

Fig. 3. The sample of different feature pairs for GOAF.

into integers in $[1, maxVal]$. The value of maxVal decides the total division that gradient orientation will be divided into. Larger maxVal will results in small or fine division of gradient orientation.

Then, using $g_1 = G(x, y)$ denote the gradient orientation value in position (x, y), we can get $g_2 = G(x + xof, y + yof)$ for a giving position offset (xof, yof). Different offsets means different distances and directions.

$$distance = \sqrt{xof^2 + yof2} \tag{1}$$

$$angle = arctan\left(yof/xof\right) \tag{2}$$

The two values can make a feature value pair $p(g_1, g_2)$. According to formula (4), this pair can be turned into a single value. So when we choose an offset (xof, yof), we can get a feature matrix F where

$$F(x, y) = f\left(G(x, y), G(x + xof, y + yof)\right) \tag{3}$$

The whole construction process of GOAF is shown in the Fig.4. When we choose an offset parameter, the gradient orientation matrix of each sample will be turned into an associated feature matrix. Then the matrixes from different offsets are normalized and aggregated together to get the final GOAF.

$$f(g_1, g_2) = \begin{cases} g_1\left(maxVal + 1\right) + g_2, g_1 \in G, g_2 \in G \\ 0, g_1 \notin G, g_2 \notin G \end{cases} \tag{4}$$

From above encoding, each feature of GOAF is also transferred into a channel feature, which can be easily fused with channel features ACF.

3.3 Local Difference of ACF (ACF-LD)

GOAF contains associated information of gradient orientation from different local regions. Then we will further exploit the context information of difference

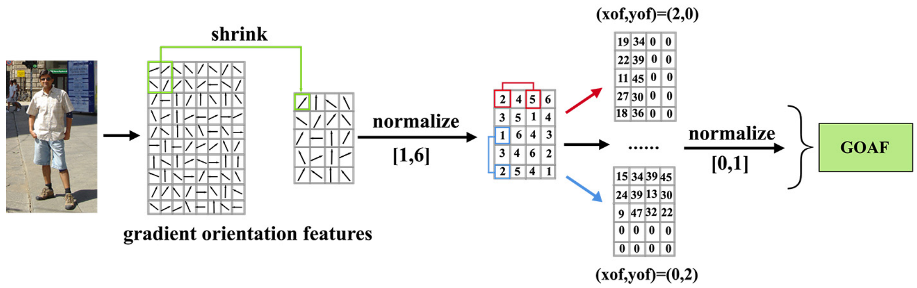

Fig. 4. The computing process of GOAF.

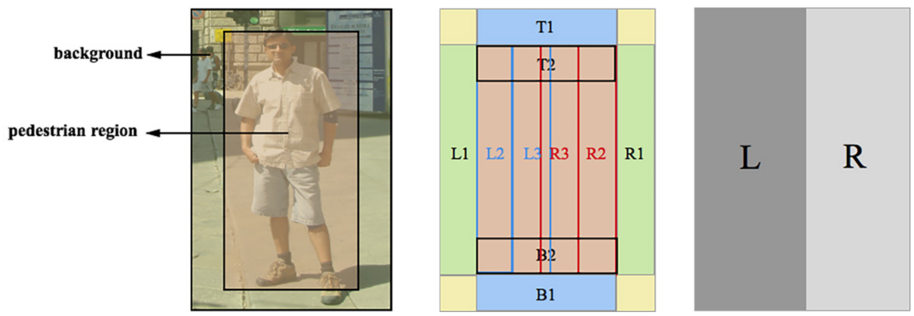

Fig. 5. The segmentation scheme of ACF-LD.

from various regions, such as pedestrian regions and neighbor background regions.

Most of the existing features focus on the pedestrian region, and do not make full use of surrounding background information. In this paper, we enlarge the detection window to contain some background regions. As Fig.5 shows the detection window or sample is divided into two parts, the target area in center and the surrounding background area. By analyzing abundant samples, we find out that the similarity be-tween human body parts is often higher than the similarity between pedestrian region and background. This contrast can help the detector distinguish the positive and negative samples. So through comparing a human with the neighborhood around it, we try to extract some useful information to describe the difference between them, further enhance the reliability of detection. Details of the construction of ACF-LD are introduced in following.

We divide the background and pedestrian areas into several sub regions as Fig.5 shows. L1, R1, T1 and B1 are used to represent the nearby background in different directions of pedestrian. L2, R2, T2 and B2 are correspondingly neighboring pedestrian areas. L3 and R3 are pedestrian areas adjacent to L2 and R2. The adjacent pedestrian and background regions, such as L1 and L2, T1 and T2, are very different, but the adjacent pedestrian regions, such as L2 and L3 are more similar to each other.

In order to capture this information effectively, the image should be expressed in reasonable feature space and the similarity between local regions should be measured appropriately. In our framework, ACF [10] features are used to represent each part. Then difference between these parts is computed, which implies the similarity between the central area and the surrounding background. In addition, pedestrian samples have some level of horizontal symmetry, so the differences of features in symmetrical positions are also computed in our method. After testing various segmentation and area combinations schemes, the above segmentation strategy is chosen as Fig.5. And there are total seven difference matrix are calculated as the following formulas show. $D_1 \sim D_4$ represent the differences between pedestrian and nearby background, D_5 and D_6 are differences within the pedestrian area, and D_7 reflects the differences of symmetrical positions in image. (Function acf is used to compute ACF matrix of each local region, and function $flip$ means to flip the matrix horizontally.)

$$D_1 = abs\left(acf\left(L_1\right) - acf\left(L_2\right)\right) \tag{5}$$

$$D_2 = abs\left(acf\left(R_1\right) - acf\left(R_2\right)\right) \tag{6}$$

$$D_3 = abs\left(acf\left(T_1\right) - acf\left(T_2\right)\right) \tag{7}$$

$$D_4 = abs\left(acf\left(B_1\right) - acf\left(B_2\right)\right) \tag{8}$$

$$D_5 = abs\left(acf\left(L_2\right) - acf\left(L_3\right)\right) \tag{9}$$

$$D_6 = abs\left(acf\left(R_2\right) - acf\left(R_3\right)\right) \tag{10}$$

$$D_7 = abs\left(acf\left(L\right) - flip\left(acf\left(R\right)\right)\right) \tag{11}$$

These above difference matrixes compose ACF-LD. In the training stage, ACF-LD is computed for every sample and each dimension of ACF-LD is fused with ACF and GOAF to form the final feature pool. To reduce computation cost, only the ACF-LD values used in the detector is needed to compute during detection process.

4 Experiments

This section introduces some experiments to evaluate the proposed local associated features and the whole detection framework. Firstly, a series of experiments are con-ducted to find out appropriate parameters related to GOAF and segmentation strategy of ACF-LD. Then the effectiveness of the two local associated features is valuated separately. Finally all features are combined together to test the detection performance and compared with Dollar's ACF framework [10] to show the effectiveness of proposed method.

Table 1. Details about the datasets.

Dataset	Type	Train			Test			M-Height
		P-img	N-img	P-num	P-img	N-img	P-num	
Inria	Photo	614	1218	1208	288	453	566	279
Caltech	Mobile	67k	61k	192k	65k	56k	155k	48

4.1 Datasets and Evaluation Methodology

All of the experiments are performed both on Inria [5] and Caltech [18] dataset. The images in Inria dataset are static pictures, which come from many different kinds of scenes. Caltech dataset contains images of video taken from a vehicle driving in an urban environment (the videos are 30Hz and we take 1 frame per second in experiments). The details of the two datasets are shown in Table.1. P-img, N-img and P-num represent the number of positive images, negative images and labeled pedestrian samples.

4.2 Performance of Local Associated Features

GOAF+ACF To exploit associated information between different local regions, GOAF associates gradient orientation values in a certain distance and direction from different regions and encodes them into a single value. As described in section 3.2, the main parameters for GOAF are the normalization factor maxVal and offset (xof, yof) between local regions. In our experiment, the original gradient orientation $(0 \sim \pi)$ is normalized into integers in $[1, maxVal]$. If $maxVal$ is too small, some gradient orientations may not be possible to distinguish from each other, and if too large, the robust associated relationships can't be described. Finally we set $maxVal$ as 6 in our experiments.

Besides, experimental results show that the value of offset can influence the performance of GOAF. Different offset means different distance and direction between two local gradient orientation values. The selection of different offset in our method is huge. To simplify the parameters selection and feature construction process, we mainly consider the horizontal and vertical offsets, which contain abundant associated information. We use xp to represent horizontal offset combinations and yp to represent vertical offset combinations. For example, $xp\&yp = [1, 2]\,\&\,[2, 3]$ means offset $(1, 0)$, $(2, 0)$, $(0, 2)$ and $(0, 3)$ are used to compute GOAF.

To test the selection of different parameters, GOAF and ACF are calculated as candidate feature pool and then a cascaded classifier with 2048 depth-two decision trees is trained by Adaboost. In our experiments, we tested over hundreds of parameter combinations. Finally, we choose offsets $[1, 2, 4, 6]\,\&\,[1, 2, 4, 6]$ to compute GOAF on Inria and $[1, 2, 4]\,\&\,[]$ on Caltech. The offsets are different on the two datasets, which means the different scene, image resolution, pedestrian posture and size would influence the selection of offsets.

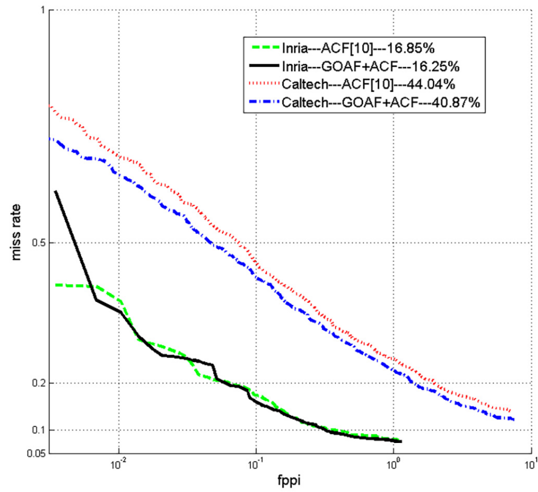

Fig. 6. Experimental results of GOAF with ACF on Caltech and Inria.

As Fig.6 shows the average miss rates are 40.87 % on Caltech based on our method, obviously better than Dollar's 44.04%. The result shows that the selected GOAF would capture useful context information hidden in the image and is an effective supplement to ACF. The miss rate 16.25% on Inria is also slightly better than Dollar's 16.85%. While the miss rate of Dollar's [10] is already very low, it's harder to further improve much.

ACF-LD+ACF For another local associated feature, ACF-LD is focus on the difference information between different regions including pedestrian or background regions. As introduced in section 3.3, the samples are divided into several sub regions and local differences of ACF are calculated between these regions. The key factors of ACF-LD are segmentation strategy and similarity measurement. We tested different segmentation schemes and found that the difference matrixes computed by sub pedestrian and background regions aligned in horizontal direction can capture more useful information. This could be because the image has some degree of symmetry in horizontal direction, a transition from background to pedestrian and from pedestrian to back-ground again. The experiments show that the partition described in section 3.3 is effective. ACF and ACF-LD features of samples are computed to train a cascaded classifier of 2048 depth-two decision trees. The miss rate of our results on Caltech (39.58%) is obviously lower than Dollar's (44.04%). The contrast between pedestrian and nearby background can help identify the targets. The results on Inria also become a litter better from 16.85% to 16.27%. ACF-LD exploits context associated

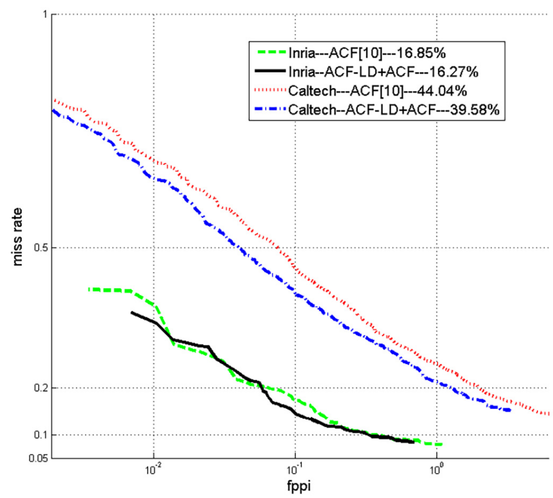

Fig. 7. Experimental results of ACF-LD with ACF on Caltech and Inria.

information of the difference between different regions of sample, which can complement with ACF to improve the performance of pedestrian detection.

GOAF+ACF-LD+ACF To further improve the robust of feature description, we fuse all the above features, such as GOAF, ACF-LD and ACF, to evaluate the overall detection performance of our framework. We use the parameters discussed in section 4.1 and 4.2. For 12864 samples on Inria, the candidate feature pool size is 15456, including 5120 ACF, 6240 ACF-LD and 4096 GOAF. And for 6432 samples on Caltech, the size of feature pool is 3234, including 1280 ACF, 1570 ACF-LD and 384 GOAF. Experiments show that the miss rate can be further reduced. We get 38.07% on Caltech, nearly 6% lower than Dollar's 44.04%. There is also 0.8% reduced on Inria from 16.85% to 16.04%. The results mean that local associated features with the channel features can robust describe pedestrian object and improve the performance of pedestrian detection.

4.3 Discussion

The above experimental results show that the proposed local associated features can effectively exploit hidden associated information and further improve the detection performance on two datasets. However, the performance improvement on Inria is very limited when compared with results on Caltech dataset. One reason is that the miss rates of Inria are already very low and hard to be further reduced. On the other side, Inria is a static image dataset including images from various kinds of scenes, such as grassland, mountain, city and country. There

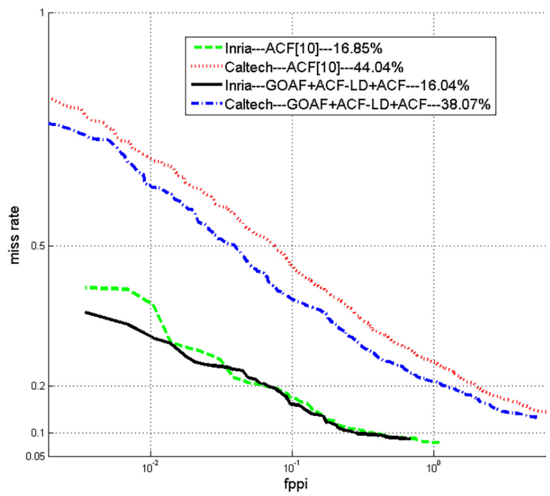

Fig. 8. Experimental results of GOAF, ACF-LF with ACF on Caltech and Inria.

are large variations in the scene of Inria dataset. Caltech are on-board video datasets with images from several urban traffic scenes. The scenes of Caltech have certain consistency and similarity.

For detection and classification task, it would be difficult to extract effective local association features for various scenes in the training process. The associated features are more suitable for scenes with stable structure information. Actually most fixed and mobile video cameras can meet the requirement, therefore, the local associated features we present can play an important role in real applications.

5 Conclusions

Context information is useful for object detection and classification, which is not made full use of in most existing pedestrian detection methods. This paper proposes two novel local associated features, GOAF and ACF-LD, to exploit the context in-formation of different local regions from pedestrian or neighbor background. These features combine information of different local regions to capture useful associated relationships. We fuse them with ACF and verify their effectiveness for pedestrian detection on different datasets. The idea and method of extracting local associated information can be applied to other algorithms conveniently.

This paper just analyses the relationship of local spatial associated information, and temporal associated information can be further studied to exploit more context information for robust pedestrian detection.

Acknowledgement. This work is supported in part by National Nature Science Foundation of China: 61202209 and in part by Beijing Natural Science Foundation: 4142051.

References

1. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition **1** (2001) 511–518
2. Ojala, M., Pietikainen, M., Maenpaa, T.: Rapid object detection using a boosted cascade of simple features. Pattern Analysis and Machine Intelligence, IEEE Transactions on **24** (2002) 971–987
3. Wu, B., Nevatia, R.: Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors. Proceedings of IEEE International Conference on Computer Vision **1** (2005) 90–97
4. Sabzmeydani, P., Mori, G.: Detecting pedestrians by learning shapelet features. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2007) 1–8
5. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition **1** (2005) 886–893
6. Wang, X., Pietikainen, T., Han, T.X., Yan, S.: An hog-lbp human detector with partial occlusion handling. Proceedings of IEEE International Conference on Computer Vision (2009) 2–29
7. Wojek, C., Schiele, B.: A performance evaluation of single and multi-feature people detection. Proceedings of DAGM Symposium on Pattern Recognition (2008) 82–91
8. Walk, S., Majer, N., Schindler, K.: New features and insights for pedestrian detection. Computer vision and pattern recognition **24** (2010) 1030–1037
9. Dollar, P., Tu, Z., Perona, P., Belongie, S.: Integral channel features. Proceedings of British Machine Vision Conference (2009) 1–11
10. Dollar, P., Appel, R., Belongie, S.: Fast feature pyramids for object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence (2014)
11. Carbonetto, P., de Freitas, N., Barnard, K.: A statistical model for general contextual object recognition. Computer Vision-ECCV (2004) 350–362
12. Hoiem, D., Efros, A.A., Hebert, M.: Putting objects in perspective. International Journal of Computer Vision **80** (2008) 3–15
13. Wolf, L., Bileschi, S.: A critical view of context. International Journal of Computer Vision **69** (2006) 251–261
14. Ramanan, D.: Using segmentation to verify object hypotheses. Computer Vision and Pattern Recognition (2007)
15. Robertson, N.M., Letham, J.: Contextual person detection in multi-modal outdoor surveillance. Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European (2012) 1930–1934
16. Schwartz, W.R., Davis, L.S., Pedrini, H.: Local response context applied to pedestrian detection. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications (2011) 181–188
17. Ding, Y., Xiao, J.: Contextual boost for pedestrian detectiont. Computer Vision and Pattern Recognition (2012) 2895–2902

18. Dollar, P., Wojek, C., Schiele, B., Perona, P.: Pedestrian detection: an evaluation of the state of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence **34** (2012) 743–761
19. Dollar, P., Belongie, S., Perona, P.: The fastest pedestrian detector in the west. Proceedings of British Machine Vision Conference (2010) 1–11
20. Tao, X., Hong, L., Yueliang, Q., Zhe, W.: A fast pedestrian detection framework based on static and dynamic information. International Conference on Multimedia and Expo (2012) 242–247
21. Kim, T.K., Cipolla, R.: Mcboost: Multiple classifiers boosting for perceptual co-clustering of images and visual features. Proceedings of IEEE Conference on Neural Information Processing Systems (2008) 841–856
22. Xu, Y., Cao, X., Qiao, H.: An efficient tree classifier ensemble-based approach for pedestrian detection. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics **41** (2011) 107–117
23. Hong, L., Tao, X., Xiangdong, W.: Robust human detection based on related hog features and cascaded adaboost and svm classifiers. Proceedings of International Conference on Multimedia Modeling (2013) 345–355
24. Dollar, P., Appel, R., Kienzle, W.: Crosstalk cascades for frame-rate pedestrian detection. Proceedings of European Conference on Computer Vision (2012) 645–659